AN INFERENCE MODEL FOR FUNCTIONAL DEPENDENCIES IN DATABASE SCHEMAS

V. Cotelea, PhD, associate professor
Academy of Economic Studies of Moldova

The design of database scheme in the third normal form [1] through synthesis method as well as quality analysis of it [2], requires to infer the functional dependencies from a set F of dependencies, which represents one part of relational schema, written $Sch(R,F)$, where the functional dependencies in F are defined over the set of attributes R.

Let F^+ be the closure of set F of functional dependencies, which represents all functional dependencies that are logically implied by F. That is, $F^+ = \{X \to Y \mid F \models X \to Y\}$ where $"\models"$ sign reads "implies logically". The $"\overset{\mathfrak{R}}{|-}"$ sign will be used to declare an inference based on a set of rules $\mathfrak{R}$. If the rule set is obvious, the $\mathfrak{R}$ designation can be avoided from above the $"|-"$ sign.

Derivation concept of a functional dependency $X \to Y$ from a set F of dependencies, assumes that exists a sequence of functional dependencies $< X_1 \to Y_1, ..., X_m \to Y_m >$, so that:

(I) $X_m \to Y_m$ is the $X \to Y$ dependency and

(II) Either $X_i \to Y_i \in F$ or either $\{X_1 \to Y_1, ..., X_{i-1} \to Y_{i-1}\} \overset{\mathfrak{R}}{|-} X_i \to Y_i$.

A set of rules $\mathfrak{R}$ is stated to be sound if $F \overset{\mathfrak{R}}{|-} X \to Y$, then $F \models X \to Y$. And the set of rules $\mathfrak{R}$ is stated to be complete if $F \models X \to Y$, then $F \overset{\mathfrak{R}}{|-} X \to Y$.

Armstrong [3] proved that the following set of rules (called axioms today) for derivation of functional dependencies is sound and complete:

- **Reflexivity**: if $Y \subseteq X$, then $X \to Y$.
- **Additivity**: if $Y \subseteq X$ and $Z \subseteq W$, then $XW \to YZ$.
- **Transitivity**: if $X \to Y$ and $Y \to Z$, then $X \to Z$.

As stated in [4] application of Armstrong axioms, from an algorithmic point of view, has a series of weaknesses. Usually a dependency can have more than one derivation sequences. These derivations are essentially equivalent, but they differ just by the order of the applied rules or by the number of times rules are applied. Besides that, derivations can contain redundant applications of rules. Actually, dependency deduction using these set of rules has an exponential complexity.

In order to eliminate the disadvantages stated in [1] a new graph-based derivation model is proposed, named *derivation tree.* Let F be a set of functional dependencies defined over the set R. The right sides of the dependencies have only one attribute. So the *derivation tree* based on F (or just *derivation tree* if the F comes from the context) is constructed according to following rules:

R1: If $A \in R$ is an attribute, then the vertex labeled with attribute A is a derivation tree.

R2: Let H be a derivation tree and A represents a suspended vertex and let in F be a dependency $B_1 ... B_m \to A$. Then the tree resulted from adding to H the vertices $B_1, ..., B_m$, as successors of the vertex A, is a derivation tree.

R3: The derivation tree based on the set F of dependencies is the tree formed only by applying rules **R1** and **R2.**

The notion of derivation tree is used to describe the deduction of functional dependencies. A functional dependency, with its derivation represented by the derivation tree, is defined by the tree's root (the right side of the dependency) and the suspended vertices (the left side of the dependency). If X is the set of suspended vertices, then the tree is called a X-tree derivation, and if the root is labeled with attribute A, then this X-tree is the derivation of the dependency $X \to A$.

In [4] the next claim is proven:

The $X \to A$ dependency is deduced from F if and only if exists an X-tree derivation with a root A formed based on the set F.

In [5] the derivation tree definition is a generalization for the case when the right side of the dependencies does not necessarily represent a single attribute. Thus, the tree is defined using the rules below:

R1: A set of isolated vertices labeled with attributes from R is a DDA-graph (*derivation directed acyclic graph*)

R2: If H is a DDA- graph and $b_1,...,b_m$ are vertices labeled with the attributes $B_1,...,B_m$, respectively, and the functional dependency $B_1,...,B_m \to CZ$ is in F, then the graph H' resulted from H by adding together the vertex c, labeled with the attribute C and the edges $(b_1,c),...,(b_m,c)$, is a DDA-graph.

R3: The resulted DDA-graph based on the dependencies in F can be obtained only by applying rules **R1** and **R2**.

In conclusion, the problem of functional dependencies membership can be solved by the enumeration of trees (graphs) and their verification (on not being a derivation tree or a DDA-graph for a respective dependency [6]). This approach is not an acceptable one due to the time consuming factor. This is why these structures hold only a theoretical nature. Besides of these, they continue to have the disadvantage of many derivations for a given dependency.

For the modeling of functional dependencies derivation in [7] is presented a structure called – *maximal derivation* (this name is taken from [8]). The construction concept is based on the algorithm which computes the closure of the set of attributes under the set of dependencies, as described in [4].

Definition 1: Let F be a set of functional dependencies over set R of attributes and let $X \subseteq R$. *Maximal derivation* of the set of attributes X under the set F of dependencies is a sequence of sets of attributes $< X_0, X_1,..., X_n >$, so that:

(I). $X_0 = X$;

(II). $X_i = X_{i-1} \cup Z$, $i = \overline{1n}$, where $Z = \cup_j W_j$ for all dependencies $V_j \to W_j \in F$ which satisfies $V_j \subseteq X_{i-1}$ and $W_j \not\subset X_{i-1}$;

(III). Nothing else from R is a member of X_i.

Before we show that maximal derivation is a powerful derivation tool for functional dependencies, two of its properties are considered.

Lemma 1. If $X \subseteq Y$ and sequences $< X_0, X_1,..., X_n >$, $< Y_0, Y_1,..., Y_m >$ are maximal derivations of the sets X and Y, respectively, under F, then for every X_i exists a set Y_j such that $X_i \subseteq Y_j$ and $j \leq i$.

Proof. The approach to prove this lemma is using the mathematical induction on i. The base case is: for $i = 0$, $X_0 = Y_0$, because $X \subseteq Y$. Let the following statement be considered as true for $i = k$: which means that $X_k \subseteq Y_p$ and $p \leq k$. Now using the induction hypothesis assumption the statement for $i = k + 1$ will be proven.

Indeed, at step $k+1$, $X_{k+1} = X_k \cup Z$, where $Z = \cup_j W_j$ for all dependencies $V_j \to W_j$ which have their right and left sides satisfying the conditions, $V_j \subseteq X_k$ and $W_j \not\subset X_k$, respectively. Based on the induction hypothesis $X_k \subseteq Y_p$ takes place. So, all the left sides of the dependencies $V_j \to W_j$ which are contained in X_k will be contained in Y_p, too. The fact that the set Y_p is larger, it can contain all the right sides W_j and then $X_{k+1} \subseteq Y_p$. If not, then in maximal derivation of the set Y under F the next $p+1$ step is executed and as a result the Y_{p+1} is obtained that will contain X_{k+1}.

This property tells us that if the set of attributes is larger, then the terms of maximal derivation converge faster and they are closer to the beginning of the maximal derivation.

Lemma 2. If $< X_0, X_1,..., X_n >$ is the maximal derivation of the set X under the set F of functional dependencies, then $X \to X_i \in F^+$, $i = \overline{0n}$.

Proof. The approach to prove this lemma is by using the mathematical induction on the number of applications of the rule (II), from the maximal derivation definition.

Consider that for the calculus of maximal derivation, the rule (II) was not applied. Then the maximal derivation consists only from one element X_0 where $X_0 = X$. From the reflexivity rule, DF1, $X \to X_0 \in F^+$ takes place.

Let on the $i-1$-th application of the rule (II), $X \to X_{i-1} \in F^+$ takes place. Next, the affirmation for the step i will be proven. Without constraining the generality, let consider that on the step i exists only one dependency $V \to W$, which satisfies $V \subseteq X_{i-1}$ and $W \not\subset X_{i-1}$. Based on the reflexivity rule: $X_{i-1} \to V \in F^+$. But from $X_{i-1} \to V \in F^+$ and $V \to W \in F^+$, applying the transitivity rule $X_{i-1} \to W \in F^+$ is obtained. From additivity rule, if the set X_i is added to the left and right side of the dependency $X_{i-1} \to W \in F^+$, the

$X_{i-1} \to X_{i-1}W \in F^+$ results. But $X_{i-1}W = X_i$, which means that $X_{i-1} \to X_i \in F^+$. That is, from $X \to X_{i-1} \in F^+$ (induction hypothesis) and $X_{i-1} \to X_i \in F^+$ the $X \to X_i \in F^+$ holds.

The property represented by this lemma states that each term of maximal derivation is functionally determined by the set of attributes on which this derivation is built.

Based on these two properties the next theorem will be proven:

Theorem 1. Let $< X_0, X_1,..., X_n >$ be the maximal derivation of the set X under the set F of functional dependencies. Then $X \to Y \in F^+$ if and only if $Y \subseteq X_n$.

Proof (Necessity). Will be proven that if $X \to Y \in F^+$, then exists a X_i in the maximal derivation $< X_0, X_1,..., X_n >$ so that $Y \subseteq X_i$. The fact that $Y \subseteq X_i$ means that $Y \subseteq X_n$ is also true. For the proof of this theorem the induction approach will be used on the number of utilized dependencies (derivation length) in the derivation of the dependency $X \to Y$ under F. The used dependency for the derivation is either in F, or it can be deduced from the reflexivity rule, or from the additivity rule, applied on the previous dependency, or by the means of the transitivity rule, applied on two previous dependencies. The last dependency of the derivation must be $X \to Y$.

Let the derivation of the dependency $X \to Y$ has length *1*, which means that it consists from the $X \to Y$ itself. There are two possible cases: either $X \to Y$ is deduced from the reflexivity rule, or $X \to Y \in F$. In the first case, $Y \subseteq X$ therefore $Y \subseteq X_0$. In the second case, the $X \to Y$ dependency will take part at the formation of the second element of the maximal derivation for the set X under F. As a result $Y \subseteq X_1$ takes place.

Now lets suppose that the affirmation is true for a derivation with a length less than k and we have to prove that the affirmation is also true for a derivation with a length equal to k. The inference rules that can be applied at this step are considered consecutively.

If for the deduction of the dependency $X \to Y$ is applied the reflexivity rule or $X \to Y \in F$, then Y behaves as for the derivations with length 1, therefore Y will be contained in X_0 and X_1 respectively.

But if $X \to Y$ results from the addition rule applied on a previous dependency $V \to W$, then exists S and T, where $T \subseteq S$ and $VS = X$ and $WT = Y$. Such that $V \to W$ has a derivation with a length less than k, then based on the induction hypothesis, there exists a set V_j in the maximal derivation, where $W \subseteq V_j$. Due to the fact that $V \subseteq X$, then from *Lemma 1* results that in the maximal derivation of X under F there is a set X_i, where $W \subseteq X_i$. From the fact that $T \subseteq S \subseteq X$ results that $T \subseteq X_0$ and $T \subseteq X_i$.

Now the last case is considered - the situation when the dependency $X \to Y$ is obtained by applying the transitivity rule on the two previous dependencies $X \to Z$ and $Z \to Y$, and both of these dependencies have a derivation with a length less than k.

Applying the induction hypothesis, for $X \to Z$ and $Z \to Y$ we have $Z \subseteq X_j$ and $Y \subseteq Z_p$ respectively. But Z_p is a term of the maximal derivation of the set Z under F. Because $Z \subseteq X_j$, from *Lemma 1* results that $Z_p \subseteq X_{j+m}$, where X_{j+m} is the $m+1$-th term of the maximal derivation of the set X_j under F, marked as $< X_{j+0}, X_{j+1},..., X_{j+m},..., X_{j+p} >$. It's obvious that the maximal derivation of the set X_j is just a subsequence of the last $n - j + 1$ elements of the maximal derivation of the set X under F. So, $Y \subseteq X_i$, where $i = j + m$.

Proof (Sufficiency). Let $< X_0, X_1,..., X_n >$ be the maximal derivation of the set X under F. From *Lemma 2* results that $X \to X_n \in F^+$. As $Y \subseteq X_n$, then by applying the reflexivity rule we have that $X_n \to Y \in F^+$. According to the transitivity rule applied on $X \to X_n \in F^+$ and $X_n \to Y \in F^+$ the $X \to Y \in F^+$ is obtained. The theorem is proved.

This theorem actually proves that applying the maximal derivation for the deduction of functional dependencies from a given set of dependencies is equivalent to applying Armstrong's axioms for the dependencies deduction process, because this theorem's proof is based only on the inference of these rules. But unlike other derivation instruments, the deduction using maximal derivation is unique, i.e. there are no two different maximal derivations for the deduction of a functional dependency from a given set of dependencies.

Due to the fact that Armstrong rules are sound and complete, the maximal derivation has the same properties.

In addition, the derivation process is not nondeterministic like in the case for the deduction using the inference rules but it's a deterministic one. The deduction algorithm complexity has a linear nature, $O(\| F \|)$. Where $\| F \|$ is the number of attributes involved in F, when duplicates are also considered.

Definition 1. Let $X \to Y \in F^+$ and $< X_0, X_1,..., X_n >$ be the maximal derivation of the set X under F. Let X_i be the first element which contains the set Y. Then the subsequence $< X_0, X_1,..., X_i >$ is considered to be the *derivation* (not necessarily the maximal one) of the functional dependency $X \to Y$ under F.

From *Theorem 1* and *Definition 1* follows

Corollary 1. $X \to Y \in F^+$ then and only then when the derivation of $X \to Y$ under F exists.

Corollary 2. If $X \to Y \in F^+$ and the dependency $V \to W \in F$ is used for computing the derivation of the $X \to Y$ under F, then $X \to V \in F^+$.

The correctness of this statement logically follows from the *Lemma 2* and the reflexivity and transitivity rules.

Conclusions: The *maximal derivation* is proposed - an inference model free of the disadvantages mentioned at the beginning of this article. Two properties of this structure are proved.

The existence of a single derivation for a given dependency will be useful for proving of different assumptions about covers of functional dependencies.

Bibliography

1. Bernstein, Philip A. *Synthesizing Third Normal Form Relations from Functional Dependencies. ACM Trans. Database Syst., V.1, N 4, p.277…298, 1976 .*

2. Maier, D. *The theory of relational database. Computer Science Press, 637 p., 1983.*

3. Armstrong, W.W. *Dependency structures of data base relationships. Information processing 74,North/Holland Pub. Co., Amsterdam, p.580…583,1974.*

4. Beeri, C., Bernstein, Philip A. *Computational problems related to the design of normal form relational database. ACM Trans. Database Syst., V.301, N 4, p.752…766, 1983*

5. Maier,. D. *Minimum cover in the relational database model. Jour. Of ACM, V.27, N 4, p 664…674, 1980.*

6. Gorbatov, A.O., Pavlov P.G.., Chetvericov, V.N. *Logicheskoe upravlenie informaczionnymi proczessami. M.: Energatomizdat, 304 p., 1984.*

7. Cotelea, Vitalie. *Baze de date relaţionale: proiectare logică*. Editura ASEM, Chişinău, 290 pag, 1997.

8. Ullman, J.D. *On Kent's "Consequences of assuming a universal relation. ACM Trans. Database Syst., V.8, N 4, p.637…643, 1983 .*

Recomandat spre publicare> 24.04.2009.